

Input file Fbh62112FL.seq; Output File 62112.trans  
Sequence length 2452

CGTGTGTGTGTCCCTGCGGCGCTAAGAAGGGGAGACTGAGGCTGAGGCTGGGGAACATCGGGCAGC	M S G	3
ATG AGC GGC		9
C G L F L R T T A A A R A C R G L V V S		23
TGC GGG CTC TTC CTG CGC ACC ACG GCT GCG GCT CGT GCC TGC CGG GGT CTG GTG GTC TCT		69
T A N R R L L R T S P P V R A F A K E L		43
ACC GCG AAC CGG CGG CTA CTG CGC ACC AGC CCG CCT GTA CGA GCT TTC GCC AAA GAG CTT		129
F L G K I K K K E V F P F P E V S Q D E		63
TTC CTA GGC AAA ATC AAG AAG AAA GAA GTT TTC CCA TTT CCA GAA GTT AGC CAA GAT GAA		189
L N E I N Q F L G P V E K F F T E E V D		83
CTT AAT GAA ATC AAT CAG TTC TTG GGA CCC GTG GAA AAA TTC TTC ACT GAA GAG GTG GAC		249
S R K I D Q E G K I P D E T L E K L K S		103
TCC CGA AAA ATT GAC CAG GAA GGG AAA ATC CCA GAT GAA ACT TTG GAG AAA TTG AAG AGC		309
L G L F G L Q V P E E Y G G L G F S N T		123
CTA GGG CTT TTT GGG CTG CAA GTC CCA GAA GAA TAT GGT GGC CTG GGC TTC TCC AAC ACC		369
M Y S R L G E I I S M D G S I T V T L A		143
ATG TAC TCA AGA CTA GGG GAG ATC ATC AGC ATG GAT GGG TCC ATC ACT GTG ACC CTG GCA		429
A H Q A I G L K G I I L A G T E E Q K A		163
GCG CAC CAG GCT ATT GGC CTC AAG GGG ATC ATC TTG GCT GGC ACT GAG GAG CAG AAA GCC		489
K Y L P K L A S G E H I A A F C L T E P		183
AAA TAC TTG CCT AAA CTG GCG TCC GGG GAG CAC ATT GCA GCC TTC TGC CTC ACG GAG CCA		549
A S G S D A A S I R S R A T L S E D K K		203
GCC AGT GGG AGC GAT GCA GCC TCA ATC CGG AGC AGA GCC ACA CTA AGT GAA GAC AAG AAG		609
H Y I L N G S K V W I T N G G L A N I F		223
CAC TAC ATC CTC AAT GGC TCC AAG GTC TGG ATT ACT AAT GGA GGA CTG GCC AAT ATT TTT		669
T V F A K T E V V D S D G S V K D K I T		243
ACT GTG TTT GCA AAG ACT GAG GTC GTT GAT TCT GAT GGA TCA GTG AAA GAC AAA ATC ACA		729
A F I V E R D F G G V T N G K P E D K L		263
GCA TTC ATA GTA GAA AGA GAC TTT GGT GGA GTC ACT AAT GGG AAA CCC GAA GAT AAA TTA		789
G I R G S N T C E V H F E N T K I P V E		283
GGC ATT CGG GGC TCC AAC ACT TGT GAA GTC CAT TTT GAA AAC ACC AAG ATA CCT GTG GAA		849
N I L G E V G D G F K V A M N I L N S G		303
AAC ATC CTT GGA GAG GTC GGA GAT GGG TTT AAG GTG GCC ATG AAC ATC CTC AAC AGC GGC		909
R F S M G S V V A G L L K R L I E M T A		323
CGG TTC AGC ATG GGC AGC GTC GTG GCT GGG CTG CTC AAG AGA TTG ATT GAA ATG ACT GCT		969
E Y A C T R K Q F N K R L S E F G L I Q		343
GAG TAC GCC TGC ACA AGG AAA CAG TTT AAC AAG AGG CTC AGT GAA TTT GGA TTG ATT CAG		1029
E K F A L M A Q K A Y V M E S M T Y L T		363
GAG AAA TTT GCA CTG ATG GCT CAG AAG GCT TAC GTC ATG GAG AGT ATG ACC TAC CTC ACA		1089
A G M L D Q P G F P D C S I E A A M V K		383
GCA GGG ATG CTG GAC CAA CCT GGC TTT CCC GAC TGC TCC ATC GAG GCA GCC ATG GTG AAG		1149
V F S S E A A W Q C V S E A L Q I L G G		403

FIGURE 1A

GTG	TTC	AGC	TCC	GAG	GCC	GCC	TGG	CAG	TGT	GTG	AGT	GAG	GCG	CTG	CAG	ATC	CTC	GGG	GGC	1209	
L	G	Y	T	R	D	Y	P	Y	E	R	I	L	R	D	T	R	I	L	L	423	
TTG	GGC	TAC	ACA	AGG	GAC	TAT	CCG	TAC	GAG	CGC	ATA	CTG	CGT	GAC	ACC	CGC	ATC	CTC	CTC	1269	
I	F	E	G	T	N	E	I	L	R	M	Y	I	A	L	T	G	L	Q	H	443	
ATC	TTC	GAG	GGA	ACC	AAT	GAG	ATT	CTC	CGG	ATG	TAC	ATC	GCC	CTG	ACG	GGT	CTG	CAG	CAT	1329	
A	G	R	I	L	T	T	R	I	H	E	L	K	Q	A	K	V	S	T	V	463	
GCC	GGC	CGC	ATC	CTG	ACT	ACC	AGG	ATC	CAT	GAG	CTT	AAA	CAG	GCC	AAA	GTG	AGC	ACA	GTC	1389	
M	D	T	V	G	R	R	L	R	D	S	L	G	R	T	V	D	L	G	L	483	
ATG	GAT	ACC	GTT	GGC	CGG	AGG	CTT	CGG	GAC	TCC	CTG	GGC	CGA	ACT	GTG	GAC	CTG	GGG	CTG	1449	
T	G	N	H	G	V	V	H	P	S	L	A	D	S	A	N	K	F	E	E	503	
ACA	GGC	AAC	CAT	GGA	GTT	GTG	CAC	CCC	AGT	CTT	GCG	GAC	AGT	GCC	AAC	AAG	TTT	GAG	GAG	1509	
N	T	Y	C	F	G	R	T	V	E	T	L	L	L	R	F	G	K	T	I	523	
AAC	ACC	TAC	TGC	TTC	GGC	CGG	ACC	GTG	GAG	ACA	CTG	CTG	CTC	CGC	TTT	GGC	AAG	ACC	ATC	1569	
M	E	E	Q	L	V	L	K	R	V	A	N	I	L	I	N	L	Y	G	M	543	
ATG	GAG	GAG	CAG	CTG	GTA	CTG	AAG	CGG	GTG	GCC	AAC	ATC	CTC	ATC	AAC	CTG	TAT	GGC	ATG	1629	
T	A	V	L	S	R	A	S	R	S	I	R	I	G	L	R	N	H	D	H	563	
ACG	GCC	GTG	CTG	TCG	CGG	GCC	AGC	CGC	TCC	ATC	CGC	ATT	GGG	CTC	CGC	AAC	CAC	GAC	CAC	1689	
E	V	L	L	A	N	T	F	C	V	E	A	Y	L	Q	N	L	F	S	L	583	
GAG	GTT	CTC	TTG	GCC	AAC	ACC	TTC	TGC	GTG	GAA	GCT	TAC	TTG	CAG	AAT	CTC	TTC	AGC	CTC	1749	
S	Q	L	D	K	Y	A	P	E	N	L	D	E	Q	I	K	K	V	S	Q	603	
TCT	CAG	CTG	GAC	AAG	TAT	GCT	CCA	GAA	AAC	CTA	GAT	GAG	CAG	ATT	AAG	AAA	GTG	TCC	CAG	1809	
Q	I	L	E	K	R	A	Y	I	C	A	H	P	L	D	R	T	C	*		622	
CAG	ATC	CTT	GAG	AAG	CGA	GCC	TAT	ATC	TGT	GCC	CAC	CCT	CTG	GAC	AGG	ACA	TGC	TGA		1866	
GGC	AGG	GAC	AGT	GT	CCCC	TGCT	ACCG	CCCC	CCCC	TACCC	ATGG	CCCC	GTG	CTGG	ATG	ACT	GT	TACT	CT	TTTTTT	CAGAA
GGT	GT	TTGGG	ATTAT	CAC	AGG	TAA	AGC	CTTT	TGTT	CCCC	GTCT	GCAC	CTGA	AGGG	TTGT	CGC	CTGG	CCTGGG	GAG	AGC	CTC
TTCC	AGG	TTTT	GAC	CTGC	AGGC	AGT	GCTCT	CTAAC	AGG	ACC	ATCAC	AGC	TTCT	GAACT	GAG	CCGG	GAG	GAG	AAT	GGA	
ATT	CTG	ACCC	CTGGA	ACTGG	CGGG	TATT	CTGG	TCA	TTGAG	GAG	ACAC	CATAG	TGGA	AACT	GGGG	CTTAT	GCTG	CTGCC			
TCC	AGG	GTGT	GAG	GTGG	TGGG	ACCTGT	GTG	CAGG	TGTGG	ATAG	CCATTT	CTG	CTCA	ACC	CAC	ATTCT	CTA	AGAA	ACA		
GCT	TGAA	AGCT	CTGT	CTGGG	TCA	TTTAA	ACTAGA	AGCAG	AGGC	ACTTAA	AAAC	ATGT	ACC	AGGA	ACC	ATTTA	ACAA				
AGA	ATATA	AAAA	TGTC	ACA	ATCT	GTGT	ACTGT	TAAAA	AAAAAAAA												

FIGURE 1B

# Protein Family / Domain Matches, HMMer version 2

Searching for complete domains in PFAM

hmmpfam - search a single seq against HMM database

HMMER 2.1.1 (Dec 1998)

Copyright (C) 1992-1998 Washington University School of Medicine

HMMER is freely distributed under the GNU General Public License (GPL).

-----

HMM file: /prod/ddm/seqanal/PFAM/pfam6.4/Pfam

Sequence file: /prod/ddm/wspace/orfanal/oa-script.26629.seq

-----

Query: 62112

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
-----	-----	-----	-----	-----
Acyl-CoA_dh_M 1	Acyl-CoA dehydrogenase, middle domain	153.0	3.5e-42	
Acyl-CoA_dh 1	Acyl-CoA dehydrogenase, C-terminal dom	152.1	9.6e-42	
Acyl-CoA_dh_N 1	Acyl-CoA dehydrogenase, N-terminal dom	73.7	4.2e-19	
Polysac_deacet 1	Polysaccharide deacetylase	-43.7	1.8	

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
-----	-----	-----	-----	-----	-----	-----	-----
Acyl-CoA_dh_N	1/1	85	177	..	29 132 .]	73.7	4.2e-19
Acyl-CoA_dh_M	1/1	179	286	..	1 106 []	153.0	3.5e-42
Acyl-CoA_dh	1/1	290	441	..	1 156 []	152.1	9.6e-42
Polysac_deacet	1/1	432	580	..	1 150 []	-43.7	1.8

Alignments of top-scoring domains:

Acyl-CoA\_dh\_N: domain 1 of 1, from 85 to 177: score 73.7, E = 4.2e-19

```

      *->RRvDksgefPalrelikaLgqlGllginvPEeyGGaGad..ylaRFm
      R++D++g+ P   e +++L  lGl+g+ vPEeyGG+G +++ ++
62112   85   RKIDQEGKIP--DETLEKLKSLGLFGLQVP EYGG LGFSntMYS--- 126

      LHAQVaalviEElarvcAstgvilsvhssLgqnpilkfGseEQKkkyLpq
      + E+   ++s  v+l++h ++g+ +i+ +G+eEQK+kyLp+
62112  127 -----RLGEIISM DGSITVTLAAHQAI GLKGIILAGTEE QKAKYL PK 168

      ltsGdliga<-*
      l+sG++i+a
62112  169 LASGEHIAA      177

```

Acyl-CoA\_dh\_M: domain 1 of 1, from 179 to 286: score 153.0, E = 3.5e-42

```

      *->AlTEPgAGSDvgSlkTtAekkeGd..dyiLNGsKmWITNGgqAdwyi
      +lTEP +GSD++S++ +A+   d+++yiLNGsK+WITNGg A++++
62112  179   CLTEPASGSDAASIRSRATLS-EDkKH YILNGSKVWITNGGLANI FT 224

      VlAvT...DpakkvpgkkgitaFlVekdtpGfsiGkKedKLG lRgSdTcE
      V+A+T+  D +   + k  itaF+Ve+d+ G++ Gk+edKLG+RgS+TcE
62112  225 VFAKTevvDSDG--SVKDKITAFIVERDFGGVTNGKPEDKLGIRGSNTCE 272

      LiFEDvrvPesniL<-*
      + FE+ ++P +niL
62112  273 VHFENTKIPVENIL      286

```

Acyl-CoA\_dh: domain 1 of 1, from 290 to 441: score 152.1, E = 9.6e-42

```

      *->GkGFkyamkeLdmeRlviAaqalGlaqgaldeAinYakqRkqFGkpl
      G+GFk+am+ L+ +R+   +++ Gl+ + ++ +++Ya  RkqF k+l
62112  290   GDGFKVAMNILNSGRFSMGSVVAGLLKRLIEMTA EYACTR KQFNKRL 336

```

FIGURE 2A

```

        adfQliQfkLAdMatkLEaarllvYraAwladr.GedAKEALptskeaam
        +f liQ+k+A Ma k ++ +++Y +A d++G + ++s eaam
62112  337 SEFGLIQEKFALMAQKAYVMESMTYLTAGMLDQpGFP-----DCSIEAAM 381

        AKlfaseaAmqvattAvQilGGvGYtkdyPveRfyRDAkitqIYEGTsEI
        +K f+seaA q +++A+QilGG GYt dyP eR +RD +i I EGT+EI
62112  382 VKVFSSEAAWQCVSEALQILGGLGYTRDYPYERILRDTRILLIFEGTNEI 431

        qrlvIaRall<-*
        r Ia + l
62112  432 LRMVIALTGL 441

Polysac_deacet: domain 1 of 1, from 432 to 580: score -43.7, E = 1.8
        *->ddksvyLTFDDGPnAapayTprlLDvLkhhkvkATFFviGsnvkdnP
        +++++LT + ++ + T+r+ + Lk+ kv + G++ +d
62112  432 LRMVIALTGLQHAG--RILTTRI-HELKQAKVSTVMDTVGRRLRD-- 473

        dlarrivkeGHeigNhtwsHPdlt.....t1
        + r v+ G gNH+ HP l+++ ++ +++++ + + + +
62112  474 -SLGRTVDLG-LTGNHGVVHPSLAdsankfeentycfgrtvetlllrfgGK 521

        taeqirdeiertneaiiqatggatptlfrpPYGewsetvlsasaklGlaa
        t +++ + r+++++i+++g t++l R+ s+s ++Gl+
62112  522 TIMEEQLVLKRVANILINLYG-MTAVLSRA-----SRSIRIGLRN 560

        vlWdvDprDWsvragadaivdavlqaa<-*
        + D v ++ v a+lq+
62112  561 H-----DHEVLLANTFCVEAYLQNL 580

```

FIGURE 2B

# Protein Family / Domain Matches, HMMer version 2

Searching for complete domains in PFAM

hmmfam - search a single seq against HMM database

HMMER 2.1.1 (Dec 1998)

Copyright (C) 1992-1998 Washington University School of Medicine

HMMER is freely distributed under the GNU General Public License (GPL).

HMM file: /prod/ddm/seqanal/PFAM/pfam5.4/Pfam  
Sequence file: /prod/ddm/wspace/orfanal/oa-script.17193.seq

Query: 62112

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
Acyl-CoA_dh	Acyl-CoA dehydrogenase	399.8	1.8e-116	1
Polysac_deacet	Polysaccharide deacetylase	-43.7	1	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
Acyl-CoA_dh	1/1	85	438	..	29 394 .]	399.8	1.8e-116
Polysac_deacet	1/1	432	580	..	1 150 []	-43.7	1

Alignments of top-scoring domains:

Acyl-CoA\_dh: domain 1 of 1, from 85 to 438: score 399.8, E = 1.8e-116

\*->RRvDksgefPlrelikaLgklGllginvPEeyGgaGad..ylaRFmL

R++D++g+ P e +++L lG1+g+ vPEeyGG+G +++ ++

62112 85 RKIDQEGKIP-DETLKLSLGLFGLQVPEEYGGGLGFSntMYS----- 126

HAQVaalviEElarvcAstgvllsvhssLgqnpilrfGseEQkkyLpql

+ E+ ++s v+l++h ++g+ i+ +G+eEQk+kyLp+l

62112 127 -----RLGEIISMDGSITVTLAAHQAIGLKGIILAGTEEQAKYLPKL 169

tsGdligafAlTEPgAGSDvgSikTtAekkeGd..dyiLNGsKmWITNGg

FIGURE 2C

62112 170 +sG++i+af+lTEP +GSD++Si+ +A+ d+++yiLNGsK+WITNGg  
 ASGEHIAAFCLTEPASGSDAASIRSRATLS-EDkkHYILNGSKVWITNGG 218

62112 219 qAdwyiVlAvT...DpakkvpqkkgitaFlVekdtpGfsiGkKedKLGIR  
 A++++V+A+T+ D + + k itaF+Ve+d+ G++ Gk+edKLG+R  
 LANIFTVFAKTeVVDSGD--SVKDKITAFIVERDFGGVTNGKPEDKLGIR 266

62112 267 gSdTcELiFEDvrvPesniLGeEGeGFkyaMktLdmeRlgiAaqalGlaq  
 gS+TcE+ FE+ ++P +niLGe G+GFk+aM+ L+ +R+ +++ G++  
 GSNTCEVHFENTKIPVENILGEVGDGFKVAMNINSGRFSMGSSVAGLLK 316

62112 317 gAldeAinYAkqRkqFGkplaefQliQfKLAdMatkLEaarllvYraAwl  
 + ++ +++YA RkqF k+l ef liQ+k+A MA k ++ +++Y +A  
 RLIEMTAEYACTRKQFNKRLSEFGLIQEKFALMAQKAYVMESMTYLTAGM 366

62112 367 adr.GedAKEALptskeAAMAKlfAseiAmkvatdAvQilGGvGYtkdyp  
 d++G + ++s eAAM+K f+se+A + +++A+QilGG GYt dyP  
 LDQpGFP-----DCSIEAAMVKVFSSEAAWQCVCSEALQILGGLGYTRDYP 411

62112 412 verfyRDAkitqIYEGTSEIQRlvIaR<-\*  
 eR +RD +i I EGT+EI r Ia  
 YERILRDTRILLIFEGTNEILRMYIAL 438

Polysac\_deacet: domain 1 of 1, from 432 to 580: score -43.7, E = 1

62112 432 \*->ddksvyLTFDDGPnAapayTprlLDvLkKhkvkATFFviGsnvkdnp  
 +++++LT + ++ + T+r+ + Lk+ kv + G++ +d  
 LRMYYIALTGLQHAG--RILTTRI-HELKQAKVSTVMDTVGRRLRD-- 473

62112 474 dlarrivkeGHeignHtwsHPdlt.....t1  
 + r v+ G gNH+ HP l+++ ++ ++++ +++ ++ + +  
 -SLGRTVDLG-LTGNHGVVHPSLAdsankfeentycfgrtvetlllrfgK 521

62112 522 taegirdeiirtneaiiqatggatptlfrpPYGewsetvlsasaklGlt  
 t +++ + r+++++i+++g t++l R+ s+s ++Gl+  
 TIMEEQVLVKRVANILINLYG-MTAVLSRA-----SRSIRIGLRN 560

62112 561 vlWdvDprDWsvragadaivdavlqaa<-\*  
 + D v ++ v a+lq+  
 H-----DHEVLLANTFCVEAYLQNL 580

FIGURE 2D

TxP expression: 30K array

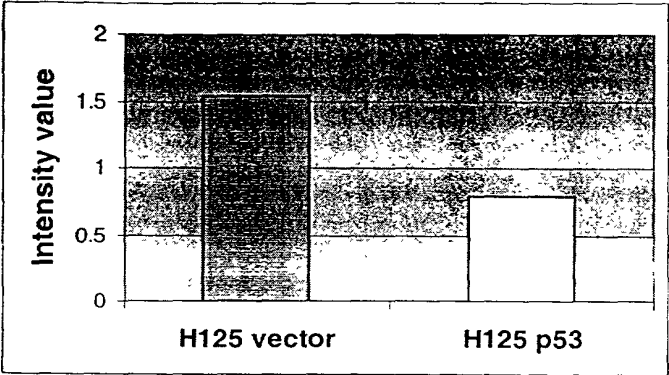


FIGURE 3A

Taqman expression

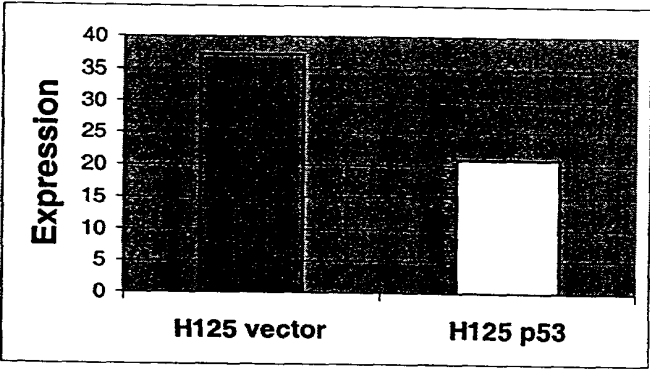


FIGURE 3B

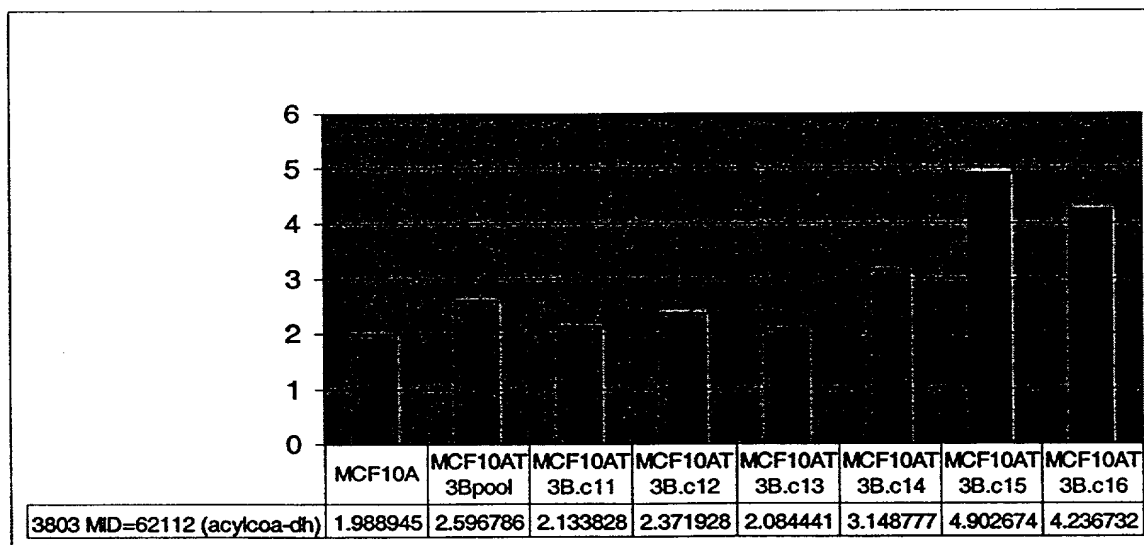


FIGURE 4

TOEBO" 92E94650

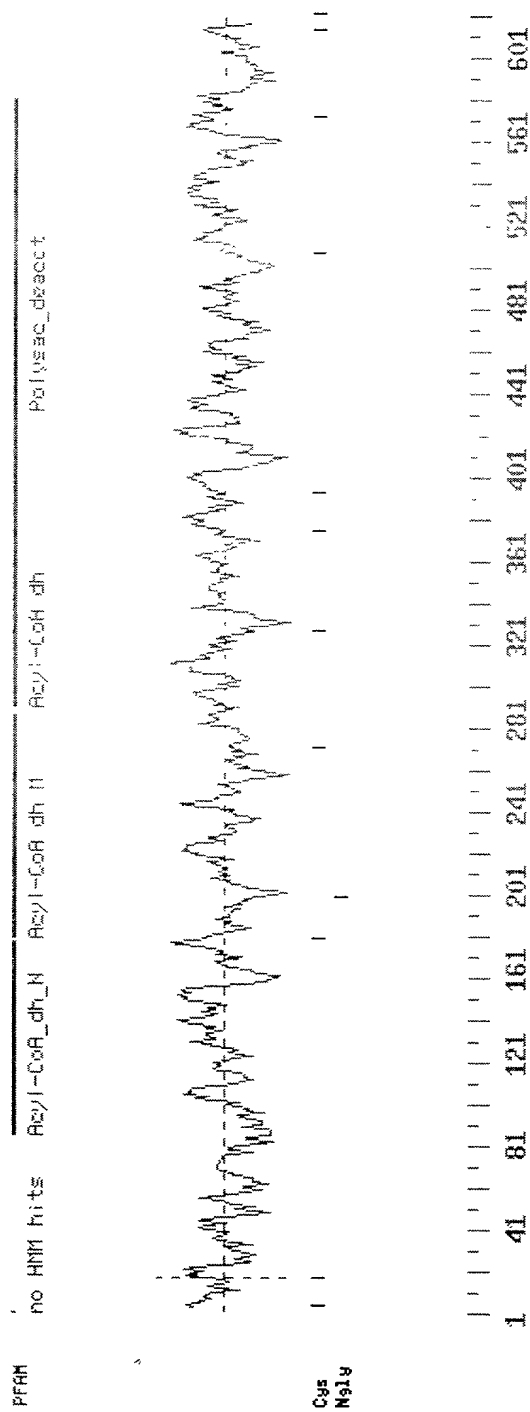


FIGURE 5